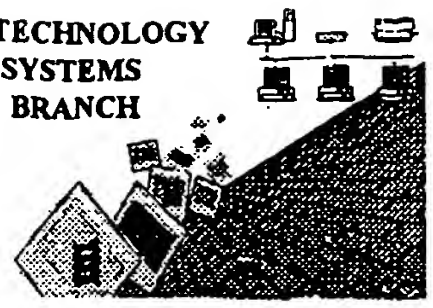
BIOTECHNOLOGY SYSTEMS BRANCH



# RAW SEQUENCE LISTING ERROR REPORT

**The Biotechnology Systems Branch of the Scientific and Technical Information Center (STIC) detected errors when processing the following computer readable form:**

Application Serial Number: 09/643,128
Source: 600
Date Processed by STIC: 11/21/03

**THE ATTACHED PRINTOUT EXPLAINS DETECTED ERRORS.**
**PLEASE FORWARD THIS INFORMATION TO THE APPLICANT BY EITHER:**

1) **INCLUDING A COPY OF THIS PRINTOUT IN YOUR NEXT COMMUNICATION TO THE APPLICANT, WITH A NOTICE TO COMPLY or,**
2) **TELEPHONING APPLICANT AND FAXING A COPY OF THIS PRINTOUT, WITH A NOTICE TO COMPLY**

**FOR CRF SUBMISSION AND PATENTIN SOFTWARE QUESTIONS, PLEASE CONTACT MARK SPENCER, TELEPHONE: 703-308-4212; FAX: 703-308-4221**
*Effective 12/13/03*: **TELEPHONE: 571-272-2510; FAX: 571-273-0221**

TO **REDUCE** ERRORED SEQUENCE LISTINGS, **PLEASE** USE THE **CHECKER VERSION 4.1 PROGRAM**, ACCESSIBLE THROUGH THE U.S. PATENT AND TRADEMARK OFFICE WEBSITE. SEE BELOW FOR ADDRESS:
**http://www.uspto.gov/web/offices/pac/checker/chkr41note.htm**

Applicants submitting genetic sequence information electronically on diskette or CD-Rom should be aware that there is a possibility that the disk/CD-Rom may have been affected by treatment given to all incoming mail.
Please consider using alternate methods of submission for the disk/CD-Rom or replacement disk/CD-Rom.
Any reply including a sequence listing in electronic form should NOT be sent to the 20231 zip code address for the United States Patent and Trademark Office, and instead should be sent via the following to the indicated addresses:

1. EFS-Bio (<**http://www.uspto.gov/ebc/efs/downloads/documents.htm**> , **EFS Submission User Manual** - ePAVE)
2. **U.S. Postal Service:** Commissioner for Patents, P.O. Box 1450, Alexandria, VA 22313-1450
3. **Hand Carry directly to (EFFECTIVE 12/01/03):**
   U.S. Patent and Trademark Office, Box Sequence, Customer Window, Lobby, Room 1B03, Crystal Plaza Two, 2011 South Clark Place, Arlington, VA 22202
4. **Federal Express, United Parcel Service, or other delivery service to:** U.S. Patent and Trademark Office, Box Sequence, Room 4B03-Mailroom, Crystal Plaza Two, 2011 South Clark Place, Arlington, VA 22202

Revised 10/08/03

# Raw Sequence Listing Error Summary

ERROR DETECTED | SUGGESTED CORRECTION | SERIAL NUMBER: 09/643,128

ATTN: NEW RULES CASES: PLEASE DISREGARD ENGLISH "ALPHA" HEADERS, WHICH WERE INSERTED BY PTO SOFTWARE

1. ____ Wrapped Nucleics / Wrapped Aminos — The number/text at the end of each line "wrapped" down to the next line. This may occur if your file was retrieved in a word processor after creating it. Please adjust your right margin to .3; this will prevent "wrapping."

2. ____ Invalid Line Length — The rules require that a line not exceed 72 characters in length. This includes white spaces.

3. ____ Misaligned Amino Numbering — The numbering under each 5th amino acid is misaligned. Do not use tab codes between numbers; use space characters, instead.

4. ____ Non-ASCII — The submitted file was not saved in ASCII(DOS) text, as required by the Sequence Rules. Please ensure your subsequent submission is saved in ASCII text.

5. ____ Variable Length — Sequence(s) ____ contain n's or Xaa's representing more than one residue. Per Sequence Rules, each n or Xaa can only represent a single residue. Please present the maximum number of each residue having variable length and indicate in the <220>-<223> section that some may be missing.

6. ____ PatentIn 2.0 "bug" — A "bug" in PatentIn version 2.0 has caused the <220>-<223> section to be missing from amino acid sequences(s) _____. Normally, PatentIn would automatically generate this section from the previously coded nucleic acid sequence. Please manually copy the relevant <220>-<223> section to the subsequent amino acid sequence. This applies to the mandatory <220>-<223> sections for Artificial or Unknown sequences.

7. ____ Skipped Sequences (OLD RULES) — Sequence(s) ____ missing. If intentional, please insert the following lines for each skipped sequence
(2) INFORMATION FOR SEQ ID NO:X: (insert SEQ ID NO where "X" is shown)
(i) SEQUENCE CHARACTERISTICS: (Do not insert any subheadings under this heading)
(xi) SEQUENCE DESCRIPTION:SEQ ID NO:X: (insert SEQ ID NO where "X" is shown)
This sequence is intentionally skipped

Please also adjust the "(ii) NUMBER OF SEQUENCES:" response to include the skipped sequences.

8. ____ Skipped Sequences (NEW RULES) — Sequence(s) _____ missing. If intentional, please insert the following lines for each skipped sequence.
<210> sequence id number
<400> sequence id number
000

9. ✓ Use of n's or Xaa's (NEW RULES) — Use of n's and/or Xaa's have been detected in the Sequence Listing.
Per 1.823 of Sequence Rules, use of <220>-<223> is MANDATORY if n's or Xaa's are present.
In <220> to <223> section, please explain location of n or Xaa, and which residue n or Xaa represents.

10. ____ Invalid <213> Response — Per 1.823 of Sequence Rules, the only valid <213> responses are: Unknown, Artificial Sequence, or scientific name (Genus/species). <220>-<223> section is required when <213> response is Unknown or is Artificial Sequence

11. ____ Use of <220> — Sequence(s) _____ missing the <220> "Feature" and associated numeric identifiers and responses.
Use of <220> to <223> is MANDATORY if <213> "Organism" response is "Artificial Sequence" or "Unknown." Please explain source of genetic material in <220> to <223> section.
(See "Federal Register," 07/01/1998, Vol. 63, No. 104, pp. 29631-32) (Sec. 1.823 of Sequence Rules)

12. ____ PatentIn 2.0 "bug" — Please do not use "Copy to Disk" function of PatentIn version 2.0. This causes a corrupted file, resulting in missing mandatory numeric identifiers and responses (as indicated on raw sequence listing). Instead, please use "File Manager" or any other manual means to copy file to floppy disk.

13. ____ Misuse of n/Xaa — "n" can only represent a single nucleotide; "Xaa" can only represent a single amino acid

IFW16

**RAW SEQUENCE LISTING** DATE: 11/21/2003
PATENT APPLICATION: **US/09/643,128** TIME: 11:01:32

Input Set : **A:\PTO.YF.txt**
Output Set: **N:\CRF4\11212003\I643128.raw**

4 <110> APPLICANT: Jifan Hu, GMR Epigenetics Corporation;
5 Andrew R. Hoffman, Stanford University
7 <120> TITLE OF INVENTION: Gene Inactivation by Targeted DNA Methylation
9 <130> FILE REFERENCE: 10853-005-999
C--> 11 **<140> CURRENT APPLICATION NUMBER: US/09/643,128**
12 <141> CURRENT FILING DATE: 2000-08-21
14 <150> PRIOR APPLICATION NUMBER: US60/96,749
15 <151> PRIOR FILING DATE: 2000-04-12
17 <160> NUMBER OF SEQ ID NOS: 56
19 <170> SOFTWARE: PatentIn For Windows v. 3

Global error.
Consult Sequence Rules.
Consult sample Sequence Listing (attached)

**ERRORED SEQUENCES**

21 <210> SEQ ID NO: 1
22 <211> LENGTH: 5
23 <212> TYPE: DNA
24 <213> ORGANISM: Artificial sequence
W--> 26 **<220> FEATURE: Hairpin** this belongs on <2237 line. Never insert a response to <2207. It is a "header" only.
W--> 27 **<221> NAME/KEY: m5C**
28 <223> OTHER INFORMATION: Phosphothioate oligonucleotide modified with 5-methylcytidine (m5C) at CpG
29 dinucleotide sequence
31 <400> SEQUENCE: 1
E--> 32 **m5cgtm5cg**
36 <210> SEQ ID NO: 2
37 <211> LENGTH: 5
38 <212> TYPE: DNA
39 <213> ORGANISM: Artificial sequence
W--> 41 **<220> FEATURE: Hairpin**
W--> 42 **<221> NAME/KEY: m5C**
43 <223> OTHER INFORMATION: Phosphothioate oligonucleotide modified with 5-methylcytidine (m5C) at CpG
44 dinucleotide sequence
46 <400> SEQUENCE: 2
E--> 47 **m5cgtm5cg**
51 <210> SEQ ID NO: 3
52 <211> LENGTH: 10
53 <212> TYPE: DNA
54 <213> ORGANISM: Artificial sequence
W--> 56 **<220> FEATURE: Hairpin**
W--> 57 **<221> NAME/KEY: m5C**
58 <223> OTHER INFORMATION: Phosphothioate oligonucleotide modified with 5-methylcytidine (m5C) at CpG

pg 1-5

delete "m5" - show nucleotide only

"n" needs explanation in <2207-<2237 section.
(see item 9 on Error Summary Sheet)

Same error

59 dinucleotide sequence
61 <400> SEQUENCE: 3

RAW SEQUENCE LISTING DATE: 11/21/2003
PATENT APPLICATION: US/09/643,128 TIME: 11:01:32

do not show daster

Input Set : A:\PTO.YF.txt
Output Set: N:\CRF4\11212003\I643128.raw

delete

E--> 62 cgacgtm5cgtm5cg
66 <210> SEQ ID NO: 4
67 <211> LENGTH: 22
68 <212> TYPE: DNA
69 <213> ORGANISM: Artificial sequence
W--> 71 <220> FEATURE: Hairpin
W--> 72 <221> NAME/KEY: m5C
73 <223> OTHER INFORMATION: Phosphothioate oligonucleotide modified with 5-methylcytidine (m5C) at CpG
74 dinucleotide sequence
76 <300> PUBLICATION INFORMATION:
77 <301> AUTHORs: Xiaoming Yao, Ji-Fan Hu, Mark Daniels, Hadas Shiran, Xiangjun Zhou, Huifan
78 Yan,
79 Hongqi Lu, Zhilan Zeng, Qingxue Wang, Tao Li, and Andrew R. Hoffman
80 <302> TITLE: A methylated oligonucleotide inhibits IGF2 expression and enhances
81 survival
82 in a model of hepatocellular carcinoma
83 <303> JOURNAL: Journal of Clinical Investigation
W--> 84 <307> DATE: in press
86 <400> SEQUENCE: 4
E--> 87 agccm5cgggm5ctgggaggagtm5cgg
91 <210> SEQ ID NO: 5
92 <211> LENGTH: 33
93 <212> TYPE: DNA
94 <213> ORGANISM: Artificial sequence
W--> 96 <220> FEATURE: Hairpin
W--> 97 <221> NAME/KEY: m5C
98 <223> OTHER INFORMATION: Phosphothioate oligonucleotide modified with 5-methylcytidine (m5C) at CpG
99 dinucleotide sequence
101 <300> PUBLICATION INFORMATION:
102 <301> AUTHORs: Xiaoming Yao, Ji-Fan Hu, Mark Daniels, Hadas Shiran, Xiangjun Zhou, Huifan
103 Yan,
104 Hongqi Lu, Zhilan Zeng, Qingxue Wang, Tao Li, and Andrew R. Hoffman
105 <302> TITLE: A methylated oligonucleotide inhibits IGF2 expression and enhances
106 survival
107 in a model of hepatocellular carcinoma
108 <303> JOURNAL: Journal of Clinical Investigation
W--> 109 <307> DATE: in press
111 <400> SEQUENCE: 6
E--> 112 cgacgtm5cgtm5cgagccm5cgggm5ctgggaggagtm5cgg
116 <210> SEQ ID NO: 6
117 <211> LENGTH: 22
118 <212> TYPE: DNA
119 <213> ORGANISM: Artificial sequence
121 <220> FEATURE:
W--> 122 <221> NAME/KEY:
123 <223> OTHER INFORMATION: Phosphothioate oligonucleotide
125 <300> PUBLICATION INFORMATION:
126 <301> AUTHORs: Xiaoming Yao, Ji-Fan Hu, Mark Daniels, Hadas Shiran, Xiangjun Zhou, Huifan

group all nucleotides into 10's, with a space between each group

22← insert cumulative base total at right margin of each line

group into 10's

33←

127 Yan,

**RAW SEQUENCE LISTING** DATE: 11/21/2003
PATENT APPLICATION: **US/09/643,128** TIME: 11:01:32

Input Set : **A:\PTO.YF.txt**
Output Set: **N:\CRF4\11212003\I643128.raw**

```
     128        Hongqi Lu, Zhilan Zeng, Qingxue Wang, Tao Li, and Andrew R. Hoffman
     129 <302> TITLE: A methylated oligonucleotide inhibits IGF2 expression and enhances
     130        survival
     131        in a model of hepatocellular carcinoma
     132 <303> JOURNAL: Journal of Clinical Investigation
W--> 133 <307> DATE: in press
     135 <400> SEQUENCE: 6
E--> 136 ggtcacggtcaggcgtagatgg
     140 <210> SEQ ID NO: 7
     141 <211> LENGTH: 21
     142 <212> TYPE: DNA
     143 <213> ORGANISM: Artificial sequence,
W--> 145 <220> FEATURE: Hairpin
W--> 146 <221> NAME/KEY: m5C
     147 <223> OTHER INFORMATION: Phosphothioate oligonucleotide modified with 5-
methylcytidine (m5C) at CpG
     148        dinucleotide sequence
     150 <400> SEQUENCE: 7
E--> 151 tm5cgctaatctcm5cgcccacm5cgg
     154 <210> SEQ ID NO: 8
     155 <211> LENGTH: 20
     156 <212> TYPE: DNA
     157 <213> ORGANISM: Artificial sequence,
W--> 159 <220> FEATURE: Hairpin
W--> 160 <221> NAME/KEY: m5C
     161 <223> OTHER INFORMATION: Phosphothioate oligonucleotide modified with 5-
methylcytidine (m5C) at CpG
     162        dinucleotide sequence
     164 <400> SEQUENCE: 8
E--> 165 acm5cggcccttttataatgm5cga
     169 <210> SEQ ID NO: 9
     170 <211> LENGTH: 20
     171 <212> TYPE: DNA
     172 <213> ORGANISM: Artificial sequence,
W--> 174 <220> FEATURE: Hairpin
W--> 175 <221> NAME/KEY: m5C
     176 <223> OTHER INFORMATION: Phosphothioate oligonucleotide modified with 5-
methylcytidine (m5C) at CpG
     177        dinucleotide sequence
     179 <400> SEQUENCE: 9
E--> 180 tcm5cgcccacm5cggcccttttat
     184 <210> SEQ ID NO: 10
     185 <211> LENGTH: 18
     186 <212> TYPE: DNA
     187 <213> ORGANISM: Artificial sequence,
W--> 189 <220> FEATURE: Hairpin
W--> 190 <221> NAME/KEY: m5C
     191 <223> OTHER INFORMATION: Phosphothioate oligonucleotide modified with 5-
methylcytidine (m5C) at CpG
     192        dinucleotide sequence
     194 <400> SEQUENCE: 10
E--> 195 cam5cgtagccm5cgagagm5ctg
     199 <210> SEQ ID NO: 11
```

group into 10's
22 ← insert
c-myc
this goes on <2237 line.
Do not insert explanation on <2137 line.
group into 10's
21 ←
c-myc
20 ←
group into 10's
10's
20 ←
HIV
10's
18 ←
FYI
IMPORTANT

<110> Smith, John; Smithgene Inc.

<120> Example of a Sequence Listing

<130> 01-00001

<140> PCT/EP98/00001
<141> 1998-12-31

<150> US 08/999,999
<151> 1997-10-15

<160> 4

<170> PatentIn version 2.0

<210> 1
<211> 389
<212> DNA
<213> Paramecium sp.

<220>
<221> CDS
<222> (279)...(389)

<300>
<301> Doe, Richard
<302> Isolation and Characterization of a Gene Encoding a Protease from Paramecium sp
<303> Journal of Genes
<304> 1
<305> 4
<306> 1-7
<307> 1988-06-11
<308> 123456
<309> 1988-06-11

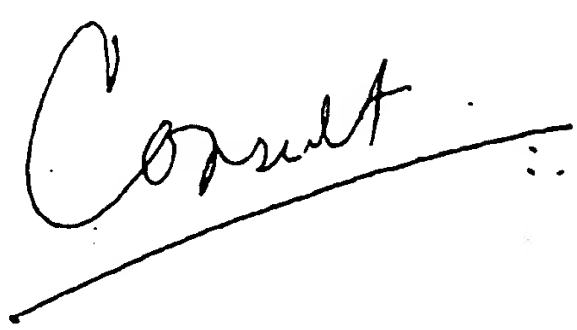
<400> 1

agctgtagtc attcctgtgt cctcttctct ctgggcttct cacccctgct atcagatctc 60

agggagagtg tcttgaccct cctctgcctt tgcagcttca caggcaggca ggcaggcagc 120

cgatgtggca attgctggca gtgccacagg ctttttagcc aggcttaggg tgggttccgc 180

cgcgggcgcgg cgggcccctct cgcgctcctc tcgcgccttt ctctcgctct cctctcgctc 240



```
ggacctgatt  aggtgagcag  gaggagggag  cagcttagc  atg gtt tca atg ttc agc  296
                                               Met Val Ser Met Phe Ser
                                                1               5

ttg tct ttc aaa tgg ccc gga ttt tgt ttg att gtt tgt ttg ttc caa  344
Leu Ser Phe Lys Trp Pro Gly Phe Cys Leu Ile Val Cys Leu Phe Gln
            10                  15                  20

tgt ccc aaa gtc ctc ccc tgt cac tca tca ctg cag ccg aat ctt      389
Cys Pro Lys Val Leu Pro Cys His Ser Ser Leu Gln Pro Asn Leu
        25                  30                  35


<210>  2
<211>  37
<212>  PRT
<213>  Paramecium sp.

<400>  2
Met Val Ser Met Phe Ser Leu Ser Phe Lys Trp Pro Gly Phe Cys Leu
 1               5                  10                  15

Phe Val Cys Leu Phe Gln Cys Pro Lys Val Leu Pro Cys His Ser Ser
            20                  25                  30

Leu Gln Pro Asn Leu
        35


<210>  3
<211>  11
<212>  PRT
<213>  Artificial Sequence

<220>
<223>  Designed peptide based on size and polarity to act as a
       linker between the alpha and beta chains of Protein XYZ.

<400>  3
Met Val Asn Leu Glu Pro Met His Thr Glu Ile
 1               5                  10


<210>  4
<400>  4
000
```

[Annex VIII follows]

identifiers and their accompanying information as shown in the following table. The numeric identifier shall be used only in the "Sequence Listing." The order and presentation of the items of information in the "Sequence Listing" shall conform to the arrangement given below. Each item of information shall begin on a new line and shall begin with the numeric identifier enclosed in angle brackets as shown. The submission of those items of information designated with an "M" is mandatory. The submission of those items of information designated with an "O" is optional. Numeric identifiers <110> through <170> shall only be set forth at the beginning of the "Sequence Listing." The following table illustrates the numeric identifiers.

| Numeric Identifier | Definition | Comments and Format | Mandatory (M) or Optional (O) |
|---|---|---|---|
| <110> | Applicant | Preferably max. of 10 names; one name per line; preferable format: Surname, Other Names and/or Initials | M |
| <120> | Title of Invention | | M |
| <130> | File Reference | Personal file reference | M, when filed prior to assignment of appl. number |
| <140> | Current Application Number | Specify as: US 07/999,999 or PCT/US96/99999 | M, if available |
| <141> | Current Filing Date | Specify as: yyyy-mm-dd | M, if available |
| <150> | Prior Application Number | Specify as: US 07/999,999 or PCT/US96/99999 | M, if applicable include priority documents under 35 USC 119 and 120 |
| <151> | Prior Application Filing Date | Specify as: yyyy-mm-dd | M, if applicable |
| <160> | Number of SEQ ID NOs | Count includes total number of SEQ ID NOs | M |
| <170> | Software | Name of software used to create the Sequence Listing | O |
| <210> | SEQ ID NO:#: | Response shall be an integer representing the SEQ ID NO shown | M |
| <211> | Length | Respond with an integer expressing the number of bases or amino acid residues | M |

| | | | |
|---|---|---|---|
| <212> | Type | Whether presented sequence molecule is DNA, RNA, or PRT (protein). If a nucleotide sequence contains both DNA and RNA fragments, the type shall be "DNA." In addition, the combined DNA/RNA molecule shall be further described in the <220> to <223> feature section. | M |
| <213> | Organism | Scientific name, i.e. Genus/species, Unknown or Artificial Sequence. In addition, the "Unknown" or "Artificial Sequence" organisms shall be further described in the <220> to <223> feature section. | M |
| <220> | Feature | Leave blank after <220>. <221-223> provide for a description of points of biological significance in the sequence. | M, under the following conditions: if "n," "Xaa," or a modified or unusual L-amino acid or modified base was used in a sequence; if ORGANISM is "Artificial Sequence" or "Unknown" or molecule is combined DNA/RNA. |
| <221> | Name/Key | Provide appropriate identifier for feature, preferably from WIPO Standard ST.25 (1998), Appendix 2, Tables 5 and 6 | M, under the following conditions: if "n," "Xaa," or a modified or unusual L-amino acid or modified base was used in a sequence |
| <222> | Location | Specify location within sequence; where appropriate state number of first and last bases/amino acids | M, under the following conditions: if "n," "Xaa," or a modified or unusual L-amino acid or modified |